Europäisches Patentamt
European Patent Office
Office européen des brevets

Bescheinigung Certificate Attestation

Die angehefteten Unterlagen stimmen mit der ursprünglich eingereichten Fassung der auf dem nächsten Blatt bezeichneten europäischen Patentanmeldung überein.

The attached documents are exact copies of the European patent application described on the following page, as originally filed.

Les documents fixés à cette attestation sont conformes à la version initialement déposée de la demande de brevet européen spécifiée à la page suivante.

Patentanmeldung Nr. Patent application No. Demande de brevet n°

03368032.3

Der Präsident des Europäischen Patentamts:
Im Auftrag

For the President of the European Patent Office

Le Président de l'Office européen des brevets
p.o.

R C van Dijk

Anmeldung Nr:
Application no.: 03368032.3
Demande no:

Anmeldetag:
Date of filing: 18.04.03
Date de dépôt:

Anmelder/Applicant(s)/Demandeur(s):

INTERNATIONAL BUSINESS MACHINES CORPORATION

Armonk, NY 10504
ETATS-UNIS D'AMERIQUE

Bezeichnung der Erfindung/Title of the invention/Titre de l'invention:
(Falls die Bezeichnung der Erfindung nicht angegeben ist, siehe Beschreibung.
If no title is shown please refer to the description.
Si aucun titre n'est indiqué se referer à la description.)

Method and system for getting data through an IP transmission network by using an optimized domain name server

In Anspruch genommene Priorität(en) / Priority(ies) claimed /Priorité(s) revendiquée(s)
Staat/Tag/Aktenzeichen/State/Date/File no./Pays/Date/Numéro de dépôt:

Internationale Patentklassifikation/International Patent Classification/
Classification internationale des brevets:

H04L29/08

Am Anmeldetag benannte Vertragstaaten/Contracting states designated at date of filing/Etats contractants désignées lors du dépôt:

AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IT LU MC NL
PT RO SE SI SK TR LI

METHOD AND SYSTEM FOR GETTING DATA THROUGH AN IP TRANSMISSION NETWORK BY USING AN OPTIMIZED DOMAIN NAME SERVER

Technical field

The present invention relates generally to the data transmission systems wherein a user can request through a transmission network pages of data stored in a content server connected to the network and relates more particularly to a method and a system for getting data through an IP transmission network by using an optimized domain name server.

Background

Today, service providers are able to provide the users of the Internet network with any kind of data which can be found in any one of the numerous content servers being accessible through the Web. In the Web context, the user has at his disposal a Web browser to access the external content servers. The access by the browser is usually done through a proxy component generally located in the service provider platform, the Web browser being enforced to go through the proxy by configuration.

The response to a request from a user may take a long time particularly in the case where the Web content server is connected through a network link with low performance or is heavy loaded. The requested document has to be transferred through the network at each request, this requiring high network performance, especially if the objects have a very

large size. One way to minimize the response time and to decrease the network loading is that the proxy used by the user be a proxy cache. The role of such a proxy cache is to intercept a request destined to a Web content server and to verify if the requested document or object is locally available in its cache, and, if it is the case, use this local copy instead of the original document.

Typically, the proxy cache stores a particular page of data only after a user has requested it. However, specified URLs (pages or more generally Web objects) may be prefetched in the proxy cache before they have effectively been requested by a user. Such a cache refreshing may have several sources such as loading specific URLs defined by the administrator, loading the most popular URLs from the previous day activity or following a specified level of HTML links on the loaded pages and caching all those linked pages.

The process of storing pages of data in a proxy cache dedicated to a user can be extended by using a set of proxy caches, all of them being able at any time to have stored the data requested by a user. But, in such a case, insofar as the same page of data may be cached in several proxy caches, this results of a network and memory load since each proxy cache must download the page in its own cache.

Solutions exist to reduce the amount of information being downloaded. A first solution, the Internet Cache Protocol (ICP) is a Web caching protocol used to exchange hints about the existence of URLs in neighbor caches. The proxy caches exchange ICP queries and replies to gather information to use in selecting the most appropriate location from which to retrieve an object. The main issue is that it implies a lot of traffic between the proxy caches without any optimization or

warranty of result. There are not any consistency between the caches and information could also be duplicated.

Another solution, the Cache Array Routing Protocol (CARP) provides seamless scaling and extreme efficiency. CARP uses a hash-based routing to provide a deterministic "request resolution path" through an array of proxies. The request resolution path, based upon a hashing of proxy array member identities and URLs, means that, for any given URL request, the browser or downstream proxy will know exactly where in the proxy array the information will be stored, if already cached from a previous request, or making a first Internet hit for delivery and caching.

Although the above solution reduces efficiently the amount of information downloaded by a Web content server to the array proxy cache, they present important drawbacks. Thus, the main drawbacks with ICP is that queries for determining the location of the cached information generate extraneous network traffic and the array of proxy caches tend to become highly redundant over a period of time insofar as proxy caches contain the same URLs of the most frequently used sites. As far as CARP is concerned, it has the drawback to need a special algorithm either in the Web browser on in a proxy to determine the localization of the requested data.

Summary of the invention

Accordingly, the main object of the invention is to achieve a method and to provide a system for determining in an efficient way whether there is a proxy cache amongst all the proxy caches being normally used in the IP network which has stored the requested data.

The invention relates therefore to a data transmission system comprising at least a data transmission network based upon the IP protocol, at least a content server able to provide data being requested by a user connected to the network, a plurality of proxies having the cache function, each of which being able to have stored the requested data and one of the proxies being the user proxy to which is addressed the request sent by the user, and a domain name server for converting the server name provided by the user to the user proxy into the IP address of the content server. The domain name server includes table means for providing the IP address of a proxy amongst the proxies being able to have stored the requested data, this table means providing the proxy IP address to the user proxy whereby the requested data can be provided to the user by the proxy storing the requested data without requesting the wanted data to the content server.

Brief description of the drawings

The above and other objects, features and advantages of the invention will be better understood by reading the following more particular description of the invention in conjunction with the accompanying drawings wherein :

- Fig. 1 is a block-diagram representing a system wherein the invention can be implemented,
- Fig. 2 is a table of ODNS according to the invention
- Fig.3 is a flow chart of the method steps implemented in an optimized domain name server according to the invention.

Detailed description of the invention

A system according to the invention is illustrated in Fig. 1. In such a system, a user 10 is connected to the Internet network 12 and can get data from a content server 14. A

plurality of proxy devices, such as proxy devices 16 or 18, are connected to the Internet network. These proxy servers are proxies having the cache function. This means that they store temporally data or HTML pages which are requested from a content server, such as server 14, and are transmitted to a user who has requested these data by the intermediary of the proxy. Note that, in the following, the proxy will be used to designate any one of these proxies having the cache function. Amongst the proxies which are connected to the Internet network, it is assumed that one of them, the proxy 16, is the proxy server to which any request from user 10 for getting data is addressed. When the proxy 16 receives such a request from the user 10, it looks up in its cache to determine whether the requested data have been stored in answer to a previous request from the user 10 or any other user. If it is the case, the requested data are immediately returned by proxy 16 to the host of user 10. It must be noted that the data which have been sent to user 10 are kept in the cache of the proxy 16.

Assuming that the requested data or HTML pages are not stored in the proxy 16, a request is sent from proxy 16 to a domain name server (DNS) 20, to resolve the IP address of the server. The DNS 20 is optimized according to the principles of the invention. Such a DNS is a server which provides the IP address of the content server corresponding to the host name of the server which is the name known by the user to designate the content server 14.

To provide the IP address of the content server, the Optimized DNS (ODNS) 20 has a table giving the IP address of the server for each server name. Each entry of the table as illustrated in Fig. 2 contains the following information :

- Server name which is the Fully Qualified Name to be translated into an IP address and which is contained in the URL sent by the user,
- A proxy IP address which is the address of a proxy in which the requested data has previously been stored,
- The IP address of the content server the name of which is the host name known by the user.

This table could be extended for administration or statistic purpose with, for example, the following fields :

- Availability indicates whether the content server is available or whether its access is refused,
- Date/time is the date and time of the previous request which was addressed to the same data,
- The identifier of the previous request.

It must be noted that there is not always an entry in the Optimized DNS table corresponding to the content server name. When such an entry exists in the table, the Optimized DNS can return the corresponding IP server address. Assuming that there is a proxy IP address corresponding to the host name contained in the request, this address is returned to proxy 16 which can send the request directly to the proxy such as the proxy device 18 which has a chance to contain the requested data without sending the request to the content server.

The steps of the method according to the invention implemented in the Optimized DNS 20 are now described in reference to Fig. 3. First of all, the ODNS is waiting for a request from the proxy associated with the user, that is proxy 16 in Fig. 1 which is called the requesting proxy in the following, (step 30). It is then determined whether the ODNS table contains an entry corresponding to the server name defined in the request (step 32). If so, it is determined whether this entry contains a proxy IP address in column 2 of the ODNS table (step 34). If

so, it is determined whether the IP address of the proxy indicated in the table is the address of the proxy which has sent the request, that is proxy 16 of Fig. 1 (step 36). If it is not the case, the IP address of the proxy mentioned in the table is returned to the requesting proxy (step 38). As already mentioned, the proxy, that is proxy 16, can then address the proxy corresponding to the proxy having to this IP address to get the requested data. Then, the process is looped back to the first step 30 of waiting for a new request.

Going back to Step 32, when the server name indicated in the request is not an entry of the table, the ODNS provides the request to another DNS of a hierarchy of DNS taking into account the tree structure of this hierarchy based upon the subnets defined in the domain name up to the root of the structure (step 40). Normally, this resolve step, which is not a part of the invention enables to get an IP address corresponding to the server name defined in the request. Then, a new entry is added to the ODNS table, such entry being the server name with the corresponding IP address of the server name (step 42).

When the new server name and its corresponding IP address have been saved in the table (step 42) or when there is an entry corresponding to the requested server name, but no proxy IP address in the ODNS table (see step 34), it is determined whether the requesting proxy (proxy 16 in Fig. 1) is a known proxy cache (step 44). A proxy is known if it is included in a list of proxies which is provided to the ODNS when this one is configured. If the requesting proxy is known, its IP address is added in the entry of the table corresponding to the server name of the request (step 46). It must be noted that, at this stage, there is always an entry with the server name of the request, which was already in the table but without any proxy

IP address in column 2 or which has been added in the table (step 42).

If the IP address of the proxy corresponding to the server name of the request is the IP address of the requesting proxy (step 36), or if it has been determined that the requesting proxy is not a known proxy (step 44), or if the IP address of the requesting proxy which was known by the ODNS has been added to the table (step 46), the IP address of the content server is returned to the requesting proxy (step 48).

Note that, in case the proxy address of which is in the entry of the table which corresponds to the server name is the requesting proxy, this means that this proxy has already sent a request to the content server. In this case, this proxy has probably been reinitialized since this request and, therefore, has lost all the data contained in its cache, and accordingly, it is necessary to send to the resquesting proxy the address of the content server..

At the initialization of the system according to the invention, all the proxies have to be declared to the ODNS so that this one has a list of proxies as mentioned above. The configuration declaration consists of the IP addresses of the proxies to be managed and the capabilities of the proxies. To start the system, all the caches are empty so that the ODNS can optimize how the proxy will be fulfilled and send the request to the right proxy.

CLAIMS

1. Data transmission system comprising at least a data transmission network (12) based upon the IP protocol, at least a content server (14) able to provide data being requested by a user (10) connected to said network, a plurality of proxies (16, 18) having the cache function, each of which being able to have stored said requested data and one of said proxies being the user proxy to which is addressed the request sent by said user, and a domain name server (20) for converting the server name provided by said user to said user proxy into the IP address of said content server;

said data transmission system being characterized in that said domain name server includes table means for providing the IP address of a proxy amongst said proxies being able to have stored said requested data, said table means providing said proxy IP address to said user proxy whereby said requested data can be provided to said user by said proxy storing said requested data without requesting said data to said content server.

2. Method for getting data in an optimized way in a data transmission system comprising at least a data transmission network (12) based upon the IP protocol, at least a content server (14) able to provide data being requested by a user (10) connected to said network, a plurality of proxies (16, 18) having the cache function each of which being able to have stored said requested data and one of said proxies being the user proxy to which is addressed the request sent by said user, and a domain

name server (20) for converting the server name provided by said user to said user proxy into the IP address of said content server;

said method being characterized by the steps of :

a) looking for an entry of a table stored in said domain name server corresponding to said server name when said user proxy has not stored the data requested by said user, and

b) determining, when there is such an entry in said table, whether said entry includes the address of a proxy amongst said plurality of proxies able to have stored said requested data.

3. Method according to claim 2, further comprising the steps of :

c) returning the proxy IP address to the user proxy if such a proxy IP address is included in said entry corresponding to said server name, and

d) sending the user request from said user proxy to the proxy IP address included in said entry.

4. Method according to claim 2, further comprising a step of determining whether the user proxy is a known proxy when said entry in said table does not include the address of a proxy, said user proxy being a known proxy when it is contained in a list of proxies provided to said domain name server at the initialization of the system.

METHOD AND SYSTEM FOR GETTING DATA THROUGH AN IP TRANSMISSION NETWORK BY USING AN OPTIMIZED DOMAIN NAME SERVER

Abstract

Data transmission system comprising at least a data transmission network (12) based upon the IP protocol, at least a content server (14) able to provide data being requested by a user (10) connected to the network, a plurality of proxies (16, 18) having the cache function, each of which being able to have stored the requested data and one of one of them being the user proxy to which is addressed the request sent by the user, and a domain name server (20) for converting the server name provided by the user to the user proxy into the IP address of the content server. The domain name server includes table means for providing the IP address of a proxy amongst the proxies being able to have stored the requested data, this table means providing the proxy IP address to the user proxy whereby the requested data can be provided to the user by the proxy storing the requested data without requesting the wanted data to the content server.

FIG. 1

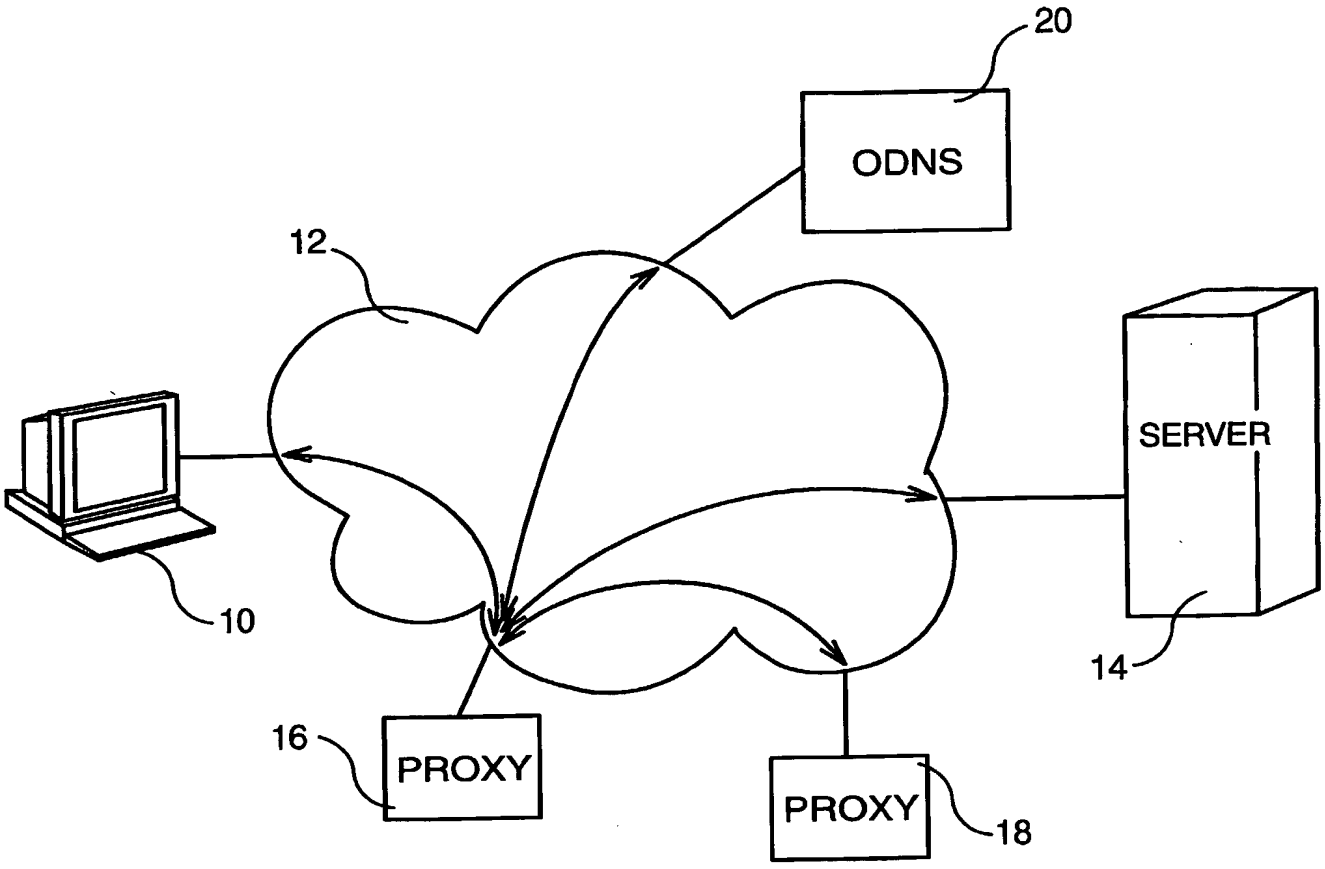

FIG. 1

SERVER NAME	PROXY IP ADDRESS	SERVER IP ADDRESS	AVAILABILITY	DATE/ TIME	REQUEST IDENTIFIER

FIG. 2

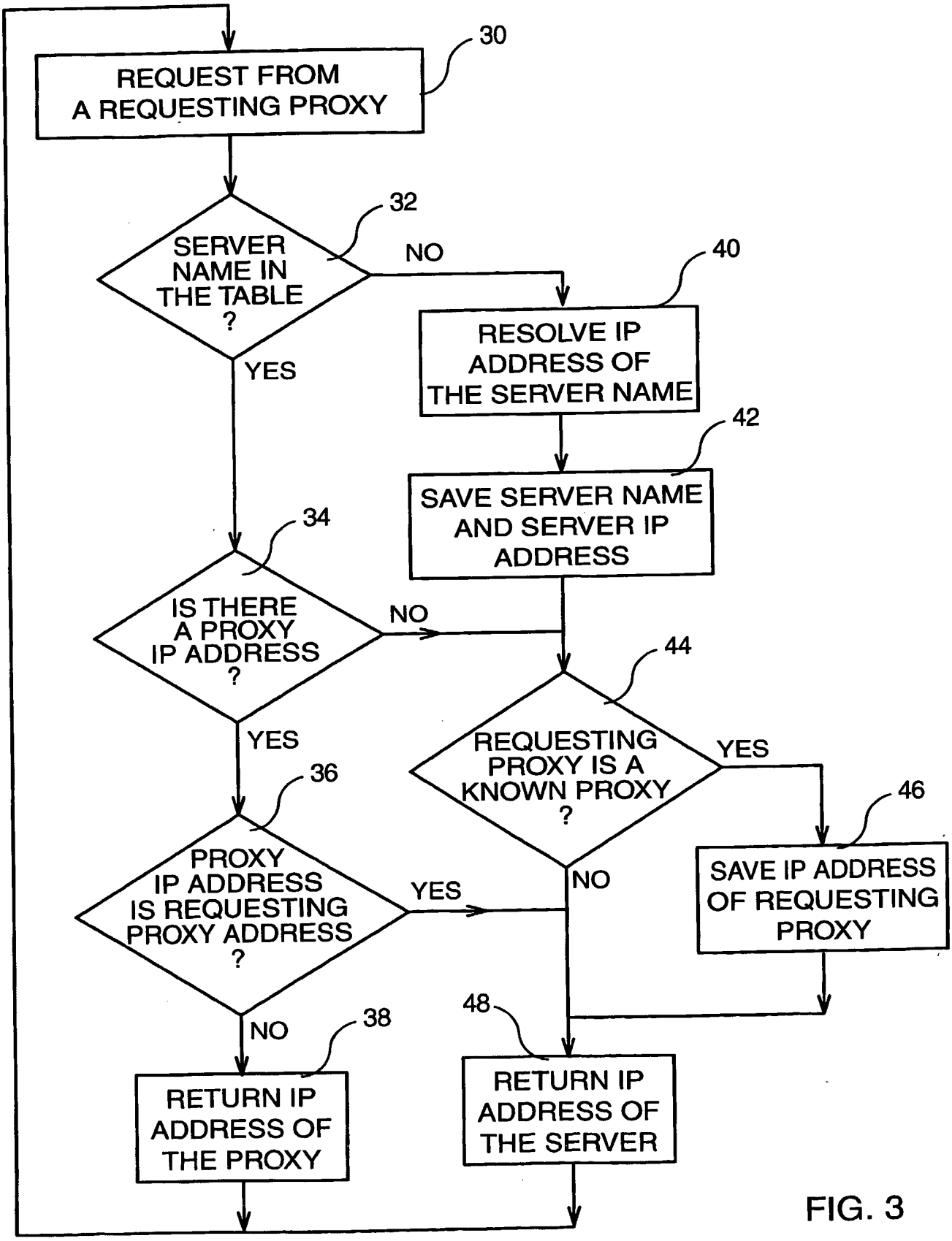
REQUEST FROM A REQUESTING PROXY
30
SERVER NAME IN THE TABLE ?
32
NO
YES
RESOLVE IP ADDRESS OF THE SERVER NAME
40
SAVE SERVER NAME AND SERVER IP ADDRESS
42
IS THERE A PROXY IP ADDRESS ?
34
NO
YES
REQUESTING PROXY IS A KNOWN PROXY ?
44
YES
NO
SAVE IP ADDRESS OF REQUESTING PROXY
46
PROXY IP ADDRESS IS REQUESTING PROXY ADDRESS ?
36
YES
NO
RETURN IP ADDRESS OF THE PROXY
38
RETURN IP ADDRESS OF THE SERVER
48

FIG. 3